

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Patent Application of:

Attila FERENCZ et al.

Application No.: Unassigned

Group Art Unit: Unassigned

Filed: February 25, 2004

Examiner: Unassigned

For: METHOD AND APPARATUS FOR SYNTHESIZING SPEECH FROM TEXT

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN  
APPLICATION IN ACCORDANCE  
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Commissioner for Patents  
PO Box 1450  
Alexandria, VA 22313-1450

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Korean Patent Application No(s). 2003-11786

Filed: February 25, 2003

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date(s) as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP



By: \_\_\_\_\_

Michael D. Stein  
Registration No. 37,240

Date: February 25, 2004

1201 New York Ave, N.W., Suite 700  
Washington, D.C. 20005  
Telephone: (202) 434-1500  
Facsimile: (202) 434-1501



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원번호 : 10-2003-0011786  
Application Number

출원년월일 : 2003년 02월 25일  
Date of Application  
FEB 25, 2003

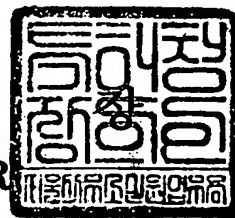
출원인 : 삼성전자주식회사  
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003      년      03      월      13      일

특      허      청

COMMISSIONER



## 【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0012
【제출일자】	2003.02.25
【국제특허분류】	G01L
【발명의 명칭】	음성 합성 방법 및 장치
【발명의 영문명칭】	Method and apparatus for text to speech synthesis
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2003-003435-0
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2003-003436-7
【발명자】	
【성명의 국문표기】	페렌츠 어틸러
【성명의 영문표기】	FERENCZ, Attila
【주소】	경기도 수원시 팔달구 영통동 955-1번지 주공아파트 140동 1506호
【국적】	RO
【발명자】	
【성명의 국문표기】	김정수
【성명의 영문표기】	KIM, Jeong Su
【주민등록번호】	640518-1347727
【우편번호】	442-374
【주소】	경기도 수원시 팔달구 매탄4동 삼성2차아파트 3동 1009호
【국적】	KR

## 【발명자】

【성명의 국문표기】

이재원

【성명의 영문표기】

LEE, Jae Won

【주민등록번호】

690301-1001549

【우편번호】

137-850

【주소】

서울특별시 서초구 방배3동 1006-1 서초ESA 3차아파트 801호

【국적】

KR

【심사청구】

청구

【취지】

특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인

이영필 (인) 대리인

이해영 (인)

## 【수수료】

【기본출원료】

20 면 29,000 원

【가산출원료】

10 면 10,000 원

【우선권주장료】

0 건 0 원

【심사청구료】

10 항 429,000 원

【합계】

468,000 원

【첨부서류】

1. 요약서·명세서(도면)\_1통

**【요약서】****【요약】**

코퍼스 기반의 음성 합성 방법이 개시된다. 코퍼스 기반으로 음성 유닛들을 접속하는 본 발명에 의한 음성 합성 방법은, (a) 접속될 상기 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하는 단계; (b) 좌측 유닛과 우측 유닛의 보간될 소정 영역을 가변적으로 결정하는 단계; (c) 좌측 유닛과 우측 유닛의 경계에 확장부를 연결하는 단계; (d) 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를 정렬하는 단계; 및 (e) 좌측 유닛 및 우측 유닛을 중첩하는 단계를 포함하는 것을 특징으로 한다. 따라서, 코퍼스 기반의 작은 크기의 DB로도, 가외 세그먼트 데이터의 유무에 따라 보간법 또는 보외법을 선택적으로 이용하여 평활화된 접속을 수행함으로써 접속 경계부의 음향적 부정합(acoustical mismatch)을 완화할 수 있고, 양호한 품질의 음성 합성을 구현할 수 있다. 본 발명은, 대용량 DB를 구비하는 시스템에서도 효과가 있으나, 상대적으로 소용량 DB를 구비하는 시스템에서 더욱 자연스럽게 만족스러운 음성 합성 출력을 제공하는 효과가 있다.

**【대표도】**

도 1

**【명세서】****【발명의 명칭】**

음성 합성 방법 및 장치{Method and apparatus for text to speech synthesis}

**【도면의 간단한 설명】**

도 1은 본 발명에 의한 코퍼스 기반의 음성 합성 방법의 바람직한 일 실시예를 설명하기 위한 플로우차트이다.

도 2는 음성 합성이 이루어져야 할 음성 유닛(speech units)들이 연속되는 구간에서의 피치 파형 및 스펙트로그램(sound spectrogram)을 나타낸다.

도 3은 도 1에 도시된 S10 단계에서, 접속이 수행되기 위하여 구분된 좌측 유닛과 우측 유닛을 나타낸다.

도 4는 도 1에 도시된 S14 단계의 바람직한 일 실시예를 설명하기 위한 플로우차트이다.

도 5는 도 1에 도시된 S14 단계에서, 가외 세그먼트 데이터를 이용한 보간법에 의하여 경계부를 확장하는 파형을 설명하기 위한 것이다.

도 6 및 도 7은 음소 경계의 한 피치를 중첩될 상대 유닛 영역에 포함된 피치 개수만큼을 반복시켜 유닛 경계에 연결하여 경계부를 확장한 것을 나타낸다.

도 6은 도 1에 도시된 S14 단계에서, 보외법에 의하여 좌측 유닛의 경계부를 확장하는 파형을 설명하기 위한 것이다.

도 7은 도 1에 도시된 S14 단계에서, 보외법에 의하여 우측 유닛의 경계부를 확장하는 파형을 설명하기 위한 것이다.

도 8은 도 1에 도시된 S16 단계에서, 우측 유닛의 보간 영역을 기준으로 좌측 유닛의 확장부 피치를 압축하여 피치 파형의 위치를 정렬하는 것을 설명하기 위한 도면이다.

도 9는 도 1에 도시된 S16 단계에서, 좌측 유닛의 보간 영역을 기준으로 우측 유닛의 확장부 피치를 신장하여 피치 파형의 위치를 정렬하는 것을 설명하기 위한 도면이다.

도 10은 도 1에 도시된 S18 단계에서, 중첩될 좌측 유닛 및 우측 유닛의 피치 트랙을 등비율적으로 보간하는 것을 설명하기 위한 도면이다.

도 11은 좌측 유닛과 우측 유닛의 중첩될 영역이 페이드인/아웃된 파형을 설명하기 위한 도면이다.

도 12는 도 11에 도시된 좌측 및 우측 유닛이 중첩된 파형을 설명하기 위한 도면이다.

도 13은 평활화 과정 없이 음소를 접속한 파형이다.

도 14는 코퍼스 기반으로 음성 유닛들을 접속하는 본 발명에 의한 음성 합성 장치를 설명하기 위한 블록도이다.

#### **【발명의 상세한 설명】**

#### **【발명의 목적】**

#### **【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<16> 본 발명은 음성 합성(Text-to-Speech Synthesis, TTS)에 관한 것으로서, 특히 코퍼스 기반(corpus-based)으로 음성 유닛들(speech units)을 접속(concatenation)하는 음성 합성 방법 및 장치에 관한 것이다.

- <17> 음성합성은 수집된 음성 데이터를 가공한 언어자료 데이터 베이스(database, 이하 "DB"라 함) 기반(corpus-based)으로 수행된다. 최근에는 음성합성이 적용되는 시스템마다 제각각의 DB의 용량에 따라, 시스템 사양을 만족하는 음성합성을 구현하는데 관심이 모아지고 있다. 즉, 대용량의 음성합성 시스템일수록 저장된 DB의 용량이 커져서 음성 데이터의 절삭 없이도 음성합성이 가능하다. 그러나, 모든 음성 합성 시스템에서 대용량 DB를 적용할 수 있는 것은 아니며, 휴대폰, 개인 휴대 정보 단말기(Personal Digital Assistant, PDA)등에는 소용량의 DB만을 적용할 수 있는 것이 현실이다. 따라서 휴대폰 등에서는 어떻게 하면 소용량의 DB만으로도 양질의 음성 합성을 구현할 수 있는지가 문제이다.
- <18> 음성 합성시에 수행되는 인접한 두 개의 음성 유닛간의 접속에 있어서는, 음향적 부정합(acoustical mismatch)을 저감하는 것이 최대의 관심사이며, 이에 대한 몇 가지의 종래기술은 다음과 같다.
- <19> 미국 특허 US5,490,234 "Waveform Blending Technique for Text-to-Speech system"은 최적 접속점을 결정하여, 단지 한 프레임에 대해서만 수행되는 평활화된 접속 즉 인접한 두 개의 피치간에 대해서만 수행되는 평활화된 접속에 관하여 개시하고 있다.
- <20> 미국 특허 출원 US 2002/0099547 "Method and Apparatus for Speech Synthesis without Prosody Modification"은 대용량의 DB와 제한된 범위에서의 중소 용량 DB에서 사용되는 음성 합성 시스템으로서, 평활화 과정 없이 대용량의 언어자료 DB를 이용한 접속에 관하여 개시하고 있다.



<21> 미국 특허 출원 US 2002/0143526 "Fast waveform synchronization for concatenation and timescale modification of speech" 는 한 피치 구간에서 수행되는 제한적인 평활화에 관한 것으로서, 정확한 피치 마킹 없이 접속될 좌 우 유닛의 경계를 조정한다.

<22> 음성 합성시에 수행되는 인접한 두 개의 유성음 음성 유닛간의 접속에 있어서는, 음향적 부정합(acoustical mismatch)을 저감하여 자연스러운 음성 합성을 구현하는 것과, 음성 합성을 구현하는 하드웨어의 자원에 따라 효율적인 음성 합성을 구현하는 것이 문제이다.

#### 【발명이 이루고자 하는 기술적 과제】

<23> 따라서 본 발명에 이루고자 하는 기술적 과제는, 음향적 고려에 기반해서 접속시 부정합을 저감하고, 언어에 독립적(language independent)으로 접속하고, 작은 DB 크기로도 양호한 음성 합성을 구현할 수 있는 음성 합성 방법을 제공하는 데 있다.

<24> 또한 본 발명이 이루고자 하는 다른 기술적 과제는, 상기 음성 합성 방법을 수행하는 음성 합성 장치를 제공하는 데 있다.

#### 【발명의 구성 및 작용】

<25> 상기한 기술적 과제를 이루기 위해, 코퍼스 기반으로 음성 유닛들을 접속하는 본 발명에 의한 음성 합성 방법은, (a) 접속될 상기 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하는 단계; (b) 상기 좌측 유닛과 상기 우측 유닛의 보간될 소정 영역을 가변적으로 결정하는 단계; (c) 상기 좌측 유닛과 상기 우측 유닛의 경계에 확장부를 연결하는 단계; (d) 상기 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를

정렬하는 단계; 및 (e) 상기 좌측 유닛 및 상기 우측 유닛을 중첩하는 단계를 포함하는 것을 특징으로 한다.

<26> 본 발명에 있어서, 상기 (c) 단계는, (c1) 상기 코퍼스내에 상기 좌측 유닛 및/또는 상기 우측 유닛에 대하여 가외 세그먼트 데이터가 존재하는가를 판단하는 단계; (c2) 상기 가외 세그먼트 데이터가 존재하는 경우에, 보간법에 의하여 경계부를 확장하는 단계; 및 (c3) 상기 가외 세그먼트 데이터가 존재하지 않는 경우에, 보외법에 의하여 경계부를 확장하는 단계;를 포함할 수 있다.

<27> 본 발명은, 상기 (d) 단계 후에 상기 중첩될 영역의 피치 트랙을 등비율적으로 보간하고 상기 (e) 단계로 진행하는 단계를 더 포함할 수도 있다.

<28> 상기한 다른 기술적 과제를 이루기 위해, 코퍼스 기반으로 음성 유닛들을 접속하는 본 발명에 의한 음성 합성 장치는, 접속될 상기 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하고, 상기 좌측 유닛과 상기 우측 유닛의 보간될 소정 영역을 가변적으로 결정하는 접속 영역 결정부; 상기 좌측 유닛과 상기 우측 유닛의 경계에 확장부를 연결하여 출력하는 경계 확장부; 상기 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를 정렬하여 출력하는 피치 파형 정렬부; 및 피치 파형 위치가 정렬된 상기 좌측 유닛 및 상기 우측 유닛을 중첩하는 유닛 중첩부를 포함하는 것을 특징으로 한다.

<29> 본 발명에 의한 음성 합성 장치에 있어서, 상기 경계 확장부는, 상기 코퍼스내에 상기 좌측 유닛 및 상기 우측 유닛에 대하여 가외 세그먼트 데이터가 존재하는가를 판단하고, 상기 가외 세그먼트 데이터가 존재하는 경우에, 보간법에 의하여 경계부를 확장하고, 상기 가외 세그먼트 데이터가 존재하지 않는 경우에, 보외법에 의하여 경계부를 확장할 수 있다.

- <30> 본 발명에 의한 음성 합성 장치는, 상기 피치 파형 정렬부로부터 피치 파형을 입력 받아, 상기 중첩될 영역의 피치 트랙을 등비율적으로 보간하여 상기 유닛 중첩부로 출력 하는 피치 트랙 보간부를 더 포함할 수 있다.
- <31> 이하, 본 발명에 의한 음성 합성 방법 및 장치를 첨부한 도면들을 참조하여 다음과 같이 상세히 설명한다.
- <32> 본 발명은 음소의 원시 데이터를 가공하여 저장한 코퍼스 기반으로 음성 유닛들을 접속하는 음성합성 방법 및 장치에 관한 것이다. 여기서 접속될 음성 유닛은 유성음 음소이다. 음소 단위로 접속을 수행함으로써 언어에 독립적으로 접속을 수행할 수 있다.
- <33> 코퍼스 기반의 음성 합성 과정은, 음성 합성을 위한 데이터 베이스를 준비하는 오프라인 처리 과정과, 입력된 문자를 준비된 데이터 베이스를 이용하여 처리하여 음성화 하는 온라인 처리 과정으로 나뉘어 진다.
- <34> 음성 합성을 위한 오프 라인 처리 과정은, 원시 데이터로부터 음성 합성을 위하여 가공된 최적 코퍼스(corpus) 선택 단계; 코퍼스 녹음(Recording) 단계; 음소 분절(Segmentation) 및 음소 정보 부여(Labeling) 단계; 데이터를 압축하여 저장하기 위한 파형 코딩(Waveform coding) 단계; 스피치 데이터 베이스화 단계; 음소-음향 파라미터 추출 단계; 음소별 정보를 데이터 베이스화 하는 단계; 데이터 절삭(pruning) 단계;를 포함한다.
- <35> 음성 합성을 위한 온라인 처리 과정은, 문자 입력 단계; 문자 전처리(preprocessing) 단계; 품사(POS, Part of speech) 분석 단계; 운율정보 생성 단계; 철자 음소 변환(Grapheme to phoneme conversion) 단계; 유닛 DB로부터 유닛을 선택하는

단계; 운율 변환(prosody conversion) 단계; 접속 및 평활화 단계; 음성 출력 단계;를 포함한다.

<36> 도 1은 본 발명에 의한 보간법 기반의 음소 접속 방법의 바람직한 일 실시예를 설명하기 위한 플로우차트로서, 접속될 음성 유닛 결정 단계(S10 단계), 보간 영역 결정 단계(S12 단계), 경계부 확장 단계(S14 단계), 피치 마커 정렬 단계(S16 단계), 피치 트랙 보간 단계(S18 단계) 및 유닛 중첩 단계(S20 단계)를 포함한다.

<37> S10 단계에서는, 접속될 음성 유닛들을 결정하여, 좌측 유닛과 우측 유닛으로 구분한다. 도 2는 음성 합성이 이루어져야 할 음성 유닛(speech units) 즉 3개의 유성음의 음소가 연속되는 구간에서의 피치 파형 및 스펙트로그램(sound spectrogram)을 나타내며, 3개의 음소간 경계부에서 파형 부정합 및 스펙트로그램의 단절을 확인할 수 있다. 음성 합성에 있어서의 평활화된 접속은, 도 1에 도시된 a-c 구간인 유성음과 유성음 사이 즉 준안정(quasi-stationary)영역에서 이루어진다. 그리고, 접속될 음성 유닛들을 결정하여 도 3과 같이 좌측 유닛과 우측 유닛으로 구분한다.

<38> S12 단계에서는, 좌측 유닛과 우측 유닛의 보간될 소정 영역을 가변적으로 결정한다. 음소간에 실제로 보간법에 의하여 접속될 보간 영역을 해당 음소의 전체 길이에 대한 비율로 결정한다. 도 2를 참조하면, 보간 영역으로 결정된 음소 길이의 40%의 위치를 알 수 있는데, 보간법에 의하여 중첩될 영역은 음성 합성의 성능 사양에 따라 가변적으로 결정될 수 있다.

<39> S14 단계는, 좌측 유닛과 우측 유닛의 경계를 기준으로 경계부를 확장하는 단계이다. S14 단계는 가외 세그먼트 데이터(extra-segmental data)를 연결하여 경계부를 확장할 수도 있고, 경계부의 소정 피치를 반복시켜 경계부를 확장할 수도 있다.

- <40> 도 4는 S14 단계의 바람직한 일 실시예를 설명하기 위한 도면으로서, 좌측 및/또는 우측 유닛의 가외 세그먼트 데이터(extra-segmental data)가 존재하는 경우와 그렇지 못한 경우로 나누어 경계부를 확장한다(S140 ~ S144 단계).
- <41> S140 단계에서는, 코퍼스(corpus)내에 좌측 유닛 및 우측 유닛의 가외 세그먼트 분절 데이터(extra-segmental data)가 존재하는가를 판단한다. 가외 세그먼트 데이터가 존재하는 경우에 보간법(Interpolation)에 의하여 평활화된 접속을 수행한다(S142 단계). 즉 경계부를 확장함에 있어서 가외 세그먼트 데이터가 존재하는 경우에는 도 5에 도시된 바와 같이, 중첩될 상대 유닛 영역의 피치 개수 만큼의 가외 세그먼트 데이터를 유닛 경계에 연결하여 경계부를 확장한다. 가외 세그먼트 데이터가 존재하지 않는 경우에 보외법(Extrapolation)에 의하여 평활화된 접속을 수행한다(S144 단계). 즉 잉여 분절 데이터가 존재하지 않는 경우에는 도 6 및 도 7에 도시된 바와 같이, 음소 경계의 한 피치를 중첩될 상대 유닛 영역에 포함된 피치 개수만큼을 반복시켜 유닛 경계에 연결하여 경계부를 확장한다.
- <42> S16 단계에서는 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치(Pitch Pulse Location)를 정렬시킨다. S16 단계는, 좌측 유닛과 우측 유닛을 중첩시키기 위한 전처리 단계로서, 대응 유닛 영역의 피치 주기에 따라 피치 파형의 위치를 정렬한다. 도 8을 참조하면, 우측 유닛의 보간 영역을 기준으로 좌측 유닛의 확장부 피치를 압축한다. 또한 도 9를 참조하면, 좌측 유닛의 보간 영역을 기준으로 우측 유닛의 확장부 피치를 신장한다.
- <43> S18 단계는 본 발명에 있어서 선택적으로 더 구비되는 단계로서, 피치 트랙을 보간하는 단계이다. S18 단계는 도 10에 도시된 바와 같이, 좌측 유닛 및 좌측 유닛의 확장

부의 피치 트랙을 등비율적으로 보간하고, 우측 유닛과 우측 유닛의 확장부의 피치 트랙을 등비율적으로 보간한다. 도 10을 참조하면, 좌측 유닛의 보간 영역의 좌측단으로부터 우측단까지 피치 간격이 균등한 비율로 감소되고 있음을 알 수 있고, 또한 우측 유닛의 보간 영역의 좌측단으로부터 우측단까지 피치 간격이 균등한 비율로 감소되고 있음을 알 수 있다.

<44> S20 단계에서는, 좌측 유닛과 우측 유닛을 중첩한다. 여기서 중첩은 페이드 인/아웃에 의하여 수행될 수 있다. 도 11은 좌측 유닛과 우측 유닛의 중첩될 영역이 페이드인/아웃된 파형을 나타낸다. 도 12는 도 11에 도시된 좌측 및 우측 유닛이 중첩된 파형을 나타낸다. 도 13은 평활화 과정 없이 음소를 접속한 파형으로서, 접속 경계부에서 급격한 파형의 변화가 있음을 알 수 있다. 이 경우 거칠고 단절된 음성으로 나타난다. 이에 비하여 도 12의 유닛간 접속 경계부를 살펴보면 급격한 파형의 변화 없이 부드럽게 접속된 것을 알 수 있다.

<45> 도 14는 코퍼스 기반으로 음성 유닛들을 접속하는 본 발명에 의한 음성 합성 장치를 설명하기 위한 블록도로서, 접속 영역 결정부(10), 경계 확장부(20), 피치 파형 정렬부(30), 유닛 중첩부(50)를 포함한다.

<46> 본 발명에 의한 음성 합성 장치는 코퍼스 기반으로 음성 유닛들을 접속한다.

<47> 접속 영역 결정부(10)는, 도 1에 도시된 S10 단계 및 S12 단계를 수행하기 위하여, 접속될 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하고, 좌측 유닛과 우측 유닛의 보간될 소정 영역을 가변적으로 결정하여 출력한다. 여기서, 접속될 음성 유닛은 유성음 음소 단위이다.

- <48>      경계 확장부(20)는, 도 1에 도시된 S14 단계를 수행하기 위하여, 좌측 유닛과 우측 유닛의 경계에 확장부를 연결하여 출력한다. 여기서 경계 확장부(20)는, 코퍼스내에 좌측 유닛 및 우측 유닛에 대하여 가외 세그먼트 데이터가 존재하는가를 판단하고, 가외 세그먼트 데이터가 존재하는 경우에, 보간법에 의하여 경계부를 확장하고, 가외 세그먼트 데이터가 존재하지 않는 경우에, 보외법에 의하여 경계부를 확장할 수 있다.
- <49>      피치 파형 정렬부(30)는, 도 1에 도시된 S16 단계를 수행하기 위하여, 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를 정렬하여 출력한다.
- <50>      유닛 중첩부(50)는, 도 1에 도시된 S20 단계를 수행하기 위하여, 피치 파형 위치가 정렬된 좌측 유닛 및 우측 유닛을 중첩하여 출력한다. 여기서 유닛 중첩부(40)는, 좌측 유닛을 페이드 아웃 하고 우측 유닛을 페이드 인 한 후에, 좌측 유닛과 우측 유닛을 중첩할 수 있다.
- <51>      본 발명에 의한 음성 합성 장치는, 도 1에 도시된 S18 단계를 수행하기 위하여, 피치 파형 정렬부(30)로부터 피치 파형을 입력받아, 중첩될 영역의 피치 트랙을 등비율적으로 보간하여 유닛 중첩부(50)로 출력하는 피치 트랙 보간부(40)를 선택적으로 더 포함할 수 있다.

#### 【발명의 효과】

- <52>      이상에서 설명한 바와 같이, 본 발명에 따른 코퍼스 기반의 음성 합성 방법에 의하면, 코퍼스 기반의 작은 크기의 DB로도, 가외 세그먼트 데이터의 유무에 따라 보간법 또는 보외법을 선택적으로 이용하여 평활화된 접속을 수행함으로써 접속 경계부의 음향적 부정합(acoustical mismatch)을 완화할 수 있고, 양호한 품질의 음성 합성을 구현할 수

있다. 본 발명은, 대용량 DB를 구비하는 시스템에서도 효과가 있으나, 상대적으로 소용량 DB를 구비하는 시스템에서 더욱 자연스럽게 만족스러운 음성 합성 출력을 제공하는 효과가 있다.

<53> 실제로 본 발명에 의해 평활화된 접속에 의해 구현된 음성과 단순 접속에 의해 구현된 음성을 비교하기 위하여, 18명의 사용자에게 3회의 설문 조사 총 54회를 시행한 결과는 다음 표 1과 같다. 표 1은 단순 접속, 가외 세그먼트 데이터를 이용한 보간법에 의한 평활화 접속, 및 보외법에 의한 평활화 접속 음성을 각각 들려주고 선호도를 1 ~ 5점으로 평가하게 한 설문 조사 결과이다.

<54> 【표 1】

	총점	평균
단순 접속	57	1.055
가외 세그먼트 데이터를 이용한 보간법에 의한 평활화 접속	233	4.314
보외법에 의한 평활화 접속	242	4.481

<55> 음소 단위로 음향적 고려에 기반해서 부정합을 저감하므로, 언어에 독립적인 (language independent) 접속을 제공한다.

<56> 본 발명은 이상에서 설명되고 도면들에 표현된 예시들에 한정되는 것은 아니다. 특히, 본 발명은 음성 합성에 있어서 유성음과 유성음 사이의 평활화된 접속을 중심으로 기재되었으나, 음성 합성 이외의 분야에서 준안정의 1차원 신호를 평활화하여 접속하려 할 때에도 본 발명이 적용될 수 있음은 명백하다.



<57>        전술한 실시 예들에 의해 가르침 받은 당업자라면, 다음의 특허 청구 범위에 기재된 본 발명의 범위 및 목적 내에서 치환, 소거, 병합, 및 단계들의 재배치 등에 의하여 전술한 실시 예들에 대해 많은 변형이 가능할 것이다.

**【특허청구범위】****【청구항 1】**

코퍼스 기반으로 음성 유닛들을 접속하는 음성 합성 방법에 있어서,

(a) 접속될 상기 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하는 단계;

(b) 상기 좌측 유닛과 상기 우측 유닛의 보간될 소정 영역을 가변적으로 결정하는 단계;

(c) 상기 좌측 유닛과 상기 우측 유닛의 경계에 확장부를 연결하는 단계;

(d) 상기 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를 정렬하는 단계; 및

(e) 상기 좌측 유닛 및 상기 우측 유닛을 중첩하는 단계를 포함하는 것을 특징으로 하는 음성 합성 방법.

**【청구항 2】**

제1항에 있어서,

상기 접속될 음성 유닛은 유성음 음소 단위인 것을 특징으로 하는 음성 합성 방법.

**【청구항 3】**

제1항에 있어서, 상기 (c) 단계는,

(c1) 상기 코퍼스내에 상기 좌측 유닛 및/또는 상기 우측 유닛에 대하여 가외 세그먼트 데이터가 존재하는가를 판단하는 단계;

(c2) 상기 가외 세그먼트 데이터가 존재하는 경우에, 보간법에 의하여 경계부를 확장하는 단계; 및

(c3) 상기 가외 세그먼트 데이터가 존재하지 않는 경우에, 보외법에 의하여 경계부를 확장하는 단계;를 포함하는 것을 특징으로 하는 음성 합성 방법.

#### 【청구항 4】

제1항에 있어서, 상기 (e) 단계는,

상기 좌측 유닛을 페이드 아웃 하고 상기 우측 유닛을 페이드 인 한 후에, 상기 좌측 유닛과 상기 우측 유닛을 중첩하는 것을 특징으로 하는 음성 합성 방법.

#### 【청구항 5】

제1항에 있어서,

(f) 상기 (d) 단계 후에 상기 중첩될 영역의 피치 트랙을 등비율적으로 보간하고 상기 (e) 단계로 진행하는 단계를 더 포함하는 것을 특징으로 하는 음성 합성 방법.

#### 【청구항 6】

코퍼스 기반으로 음성 유닛들을 접속하는 음성 합성 장치에 있어서,

접속될 상기 음성 유닛들을 결정하여 좌측 유닛과 우측 유닛으로 구분하고, 상기 좌측 유닛과 상기 우측 유닛의 보간될 소정 영역을 가변적으로 결정하는 접속 영역 결정부;

상기 좌측 유닛과 상기 우측 유닛의 경계에 확장부를 연결하여 출력하는 경계 확장부;

상기 확장부에 대하여 대응 유닛을 기준으로 피치 파형 위치를 정렬하여 출력하는  
피치 파형 정렬부; 및

피치 파형 위치가 정렬된 상기 좌측 유닛 및 상기 우측 유닛을 중첩하는 유닛 중첩  
부를 포함하는 것을 특징으로 하는 음성 합성 장치.

**【청구항 7】**

제6항에 있어서,

상기 접속될 음성 유닛은 유성음 음소 단위인 것을 특징으로 하는 음성 합성 장치.

**【청구항 8】**

제6항에 있어서, 상기 경계 확장부는,

상기 코퍼스내에 상기 좌측 유닛 및 상기 우측 유닛에 대하여 가외 세그먼트 데이  
터가 존재하는가를 판단하고, 상기 가외 세그먼트 데이터가 존재하는 경우에, 보간법에  
의하여 경계부를 확장하고, 상기 가외 세그먼트 데이터가 존재하지 않는 경우에, 보외법  
에 의하여 경계부를 확장하는 것을 특징으로 하는 음성 합성 장치.

**【청구항 9】**

제6항에 있어서, 상기 유닛 중첩부는,

상기 좌측 유닛을 페이드 아웃 하고 상기 우측 유닛을 페이드 인 한 후에, 상기 좌  
측 유닛과 상기 우측 유닛을 중첩하는 것을 특징으로 하는 음성 합성 장치.

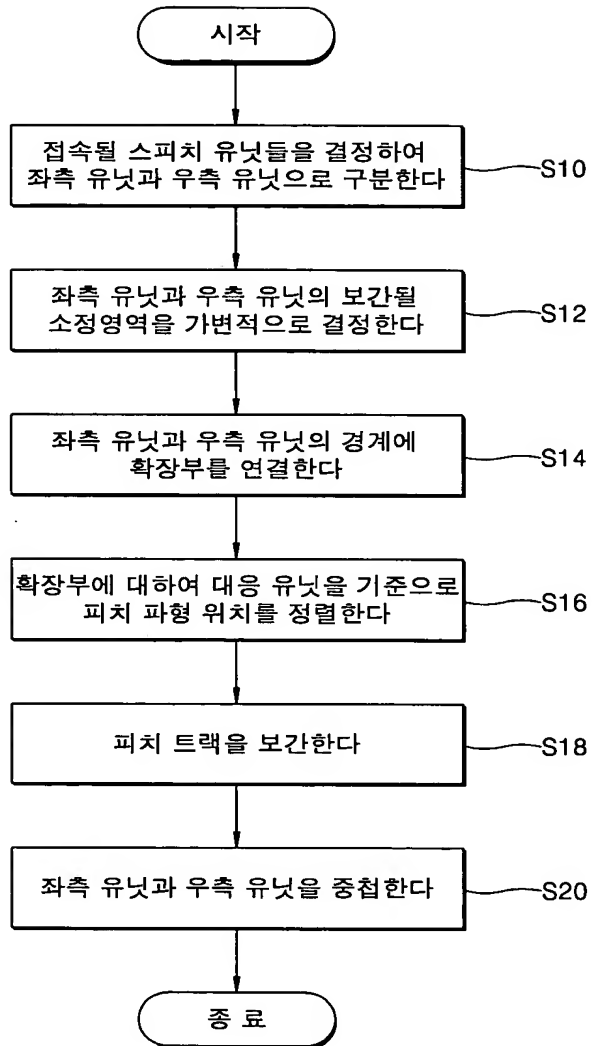
**【청구항 10】**

제6항에 있어서,

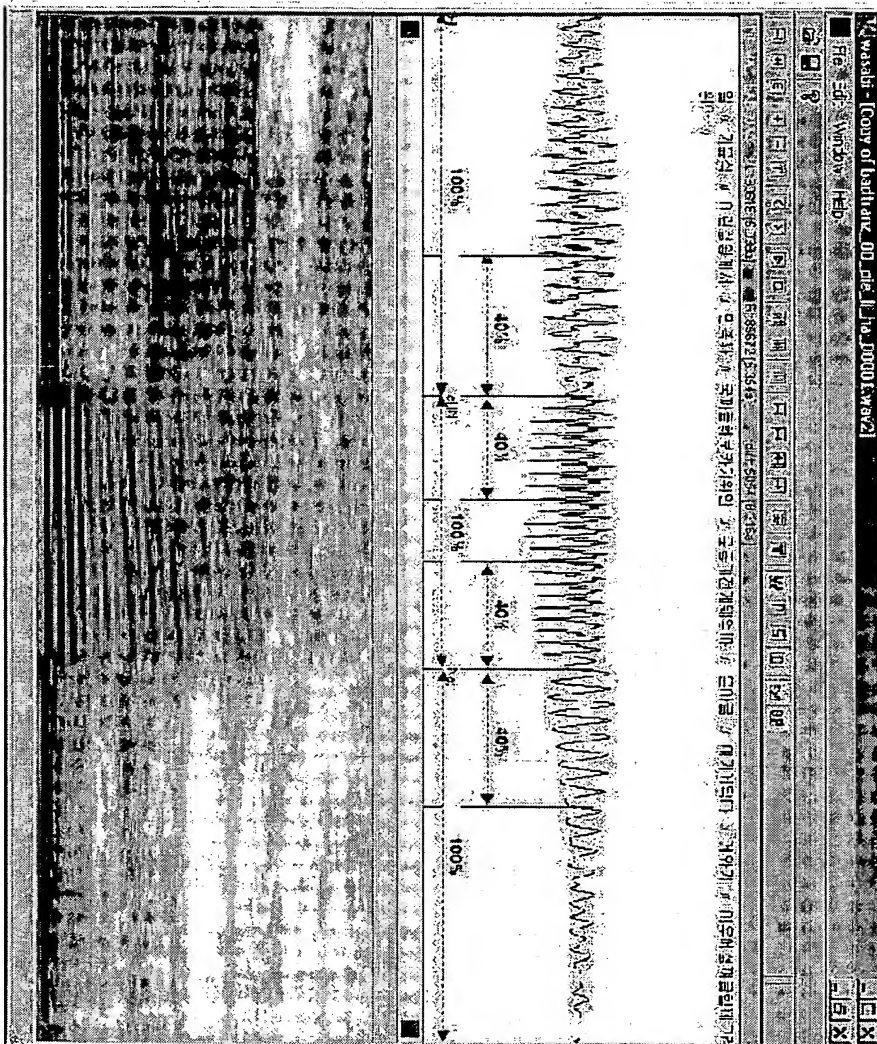
상기 피치 파형 정렬부로부터 피치 파형을 입력받아, 상기 중첩될 영역의 피치 트랙을 등비율적으로 보간하여 상기 유닛 중첩부로 출력하는 피치 트랙 보간부를 더 포함하는 것을 특징으로 하는 음성 합성 장치.

## 【도면】

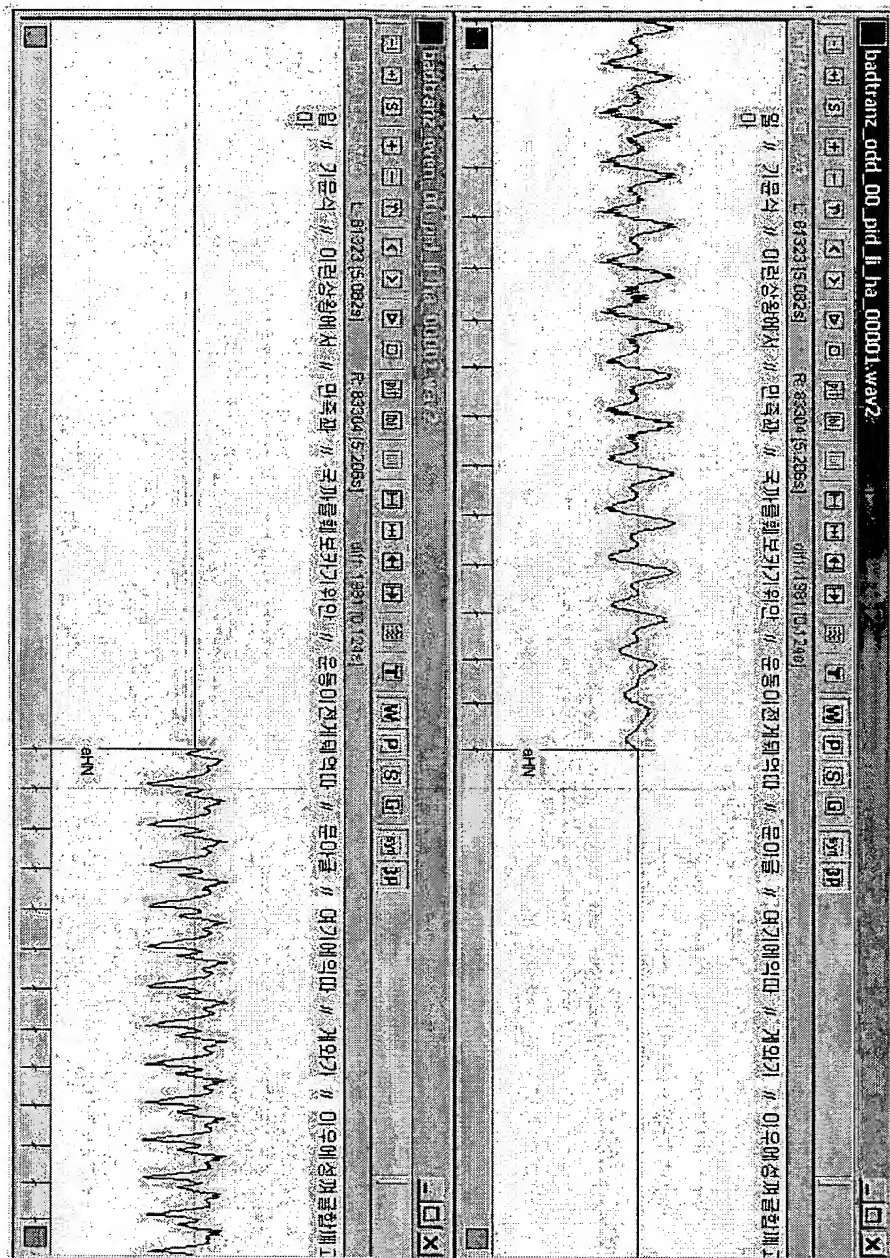
【도 1】



【도 2】

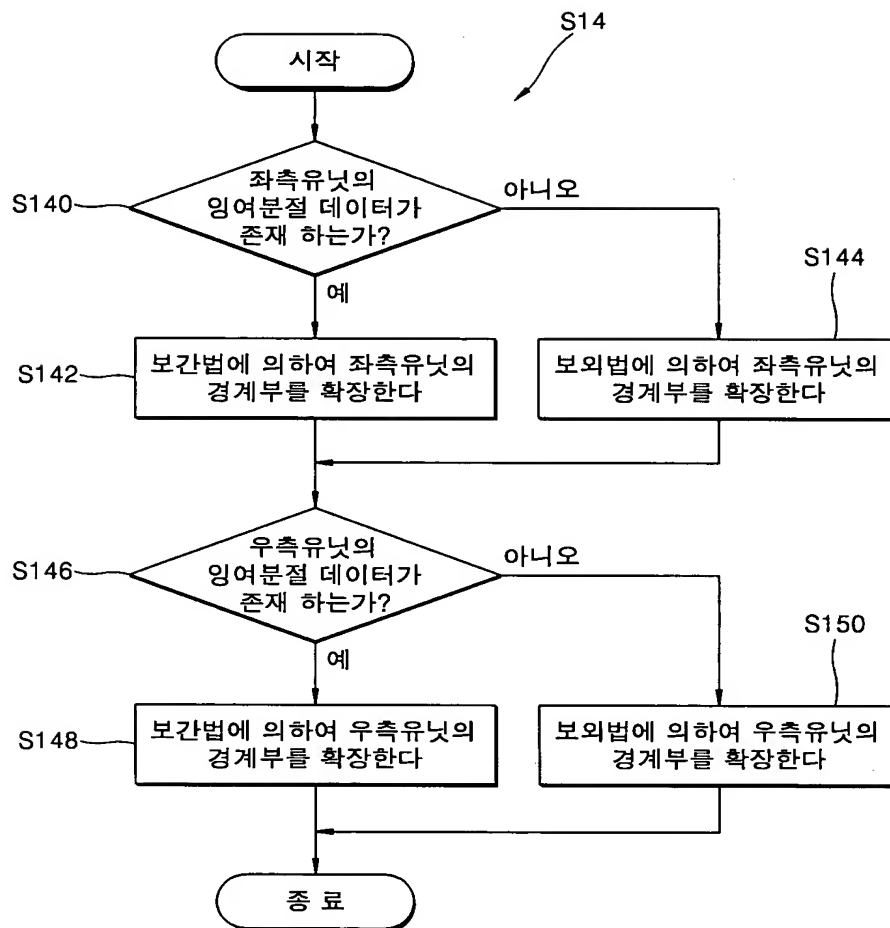


【도 3】

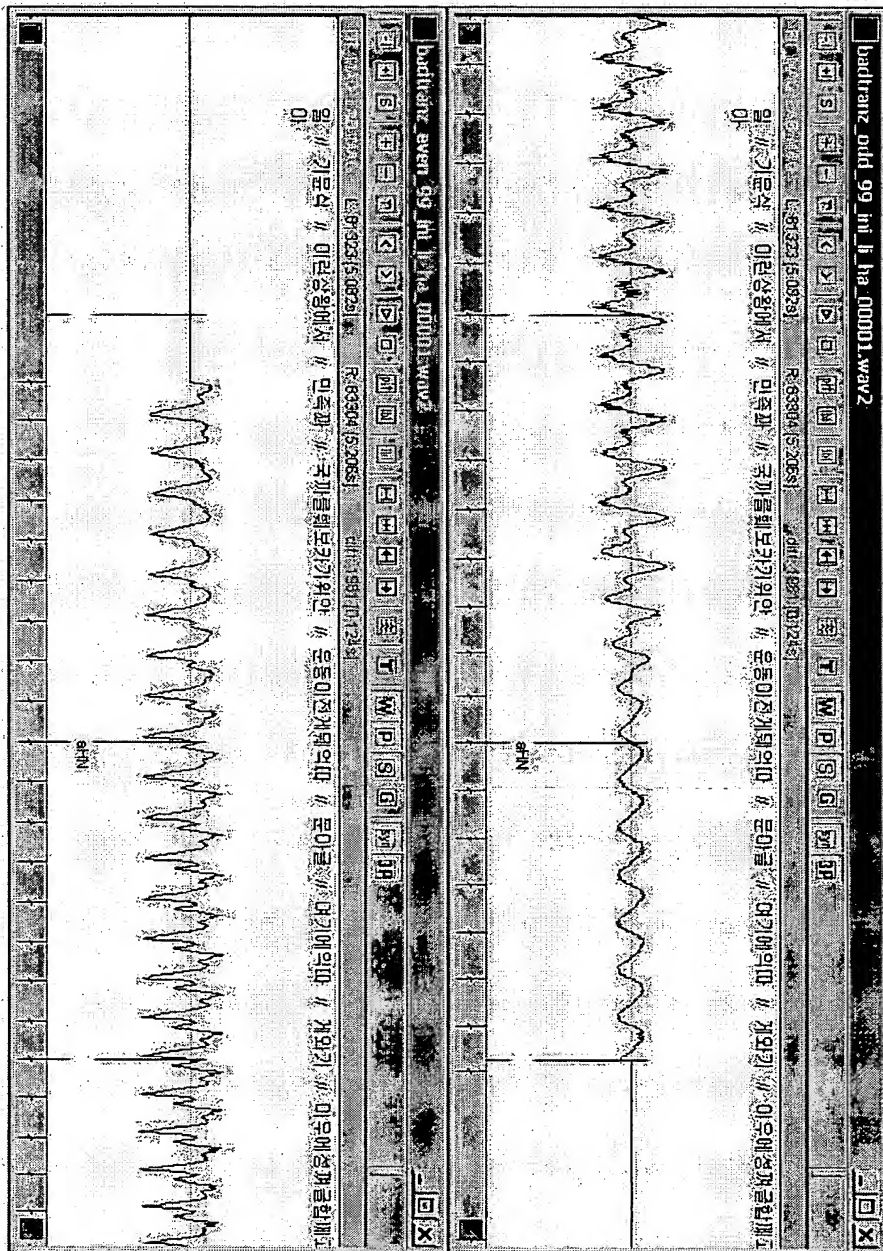




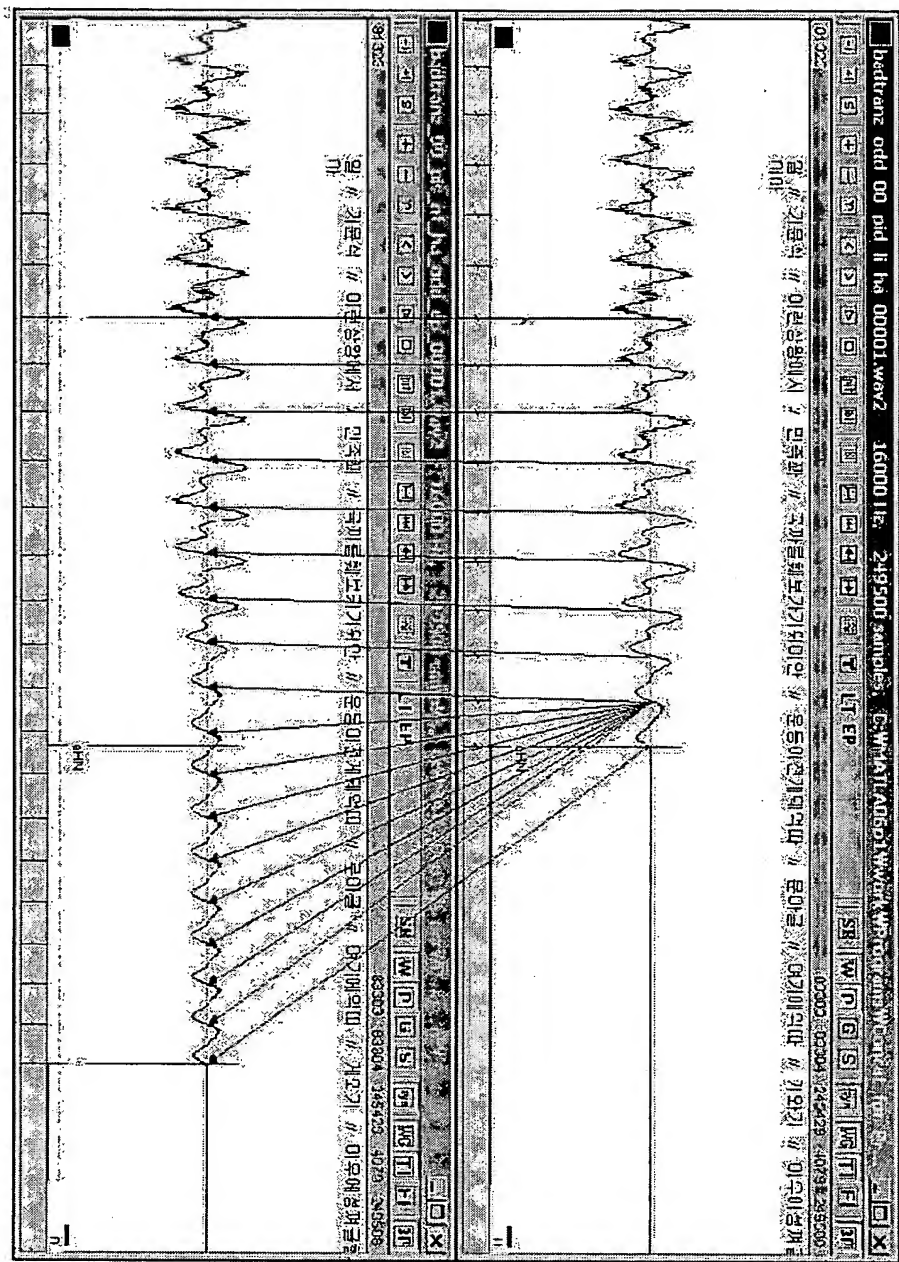
【도 4】



【도 5】



【도 6】



The figure displays a seismic waveform plot with multiple traces. The x-axis represents time in seconds, ranging from 0 to 1000. The y-axis represents amplitude. A prominent peak is visible around 1000 seconds. Below the plot is a table with columns for time, amplitude, and other parameters.

Time (s)	Amplitude	Other Parameters
0	0.00	0.00
100	0.10	0.10
200	0.20	0.20
300	0.30	0.30
400	0.40	0.40
500	0.50	0.50
600	0.60	0.60
700	0.70	0.70
800	0.80	0.80
900	0.90	0.90
1000	1.00	1.00
1100	0.90	0.90
1200	0.80	0.80
1300	0.70	0.70
1400	0.60	0.60
1500	0.50	0.50
1600	0.40	0.40
1700	0.30	0.30
1800	0.20	0.20
1900	0.10	0.10
2000	0.00	0.00

[illegible]

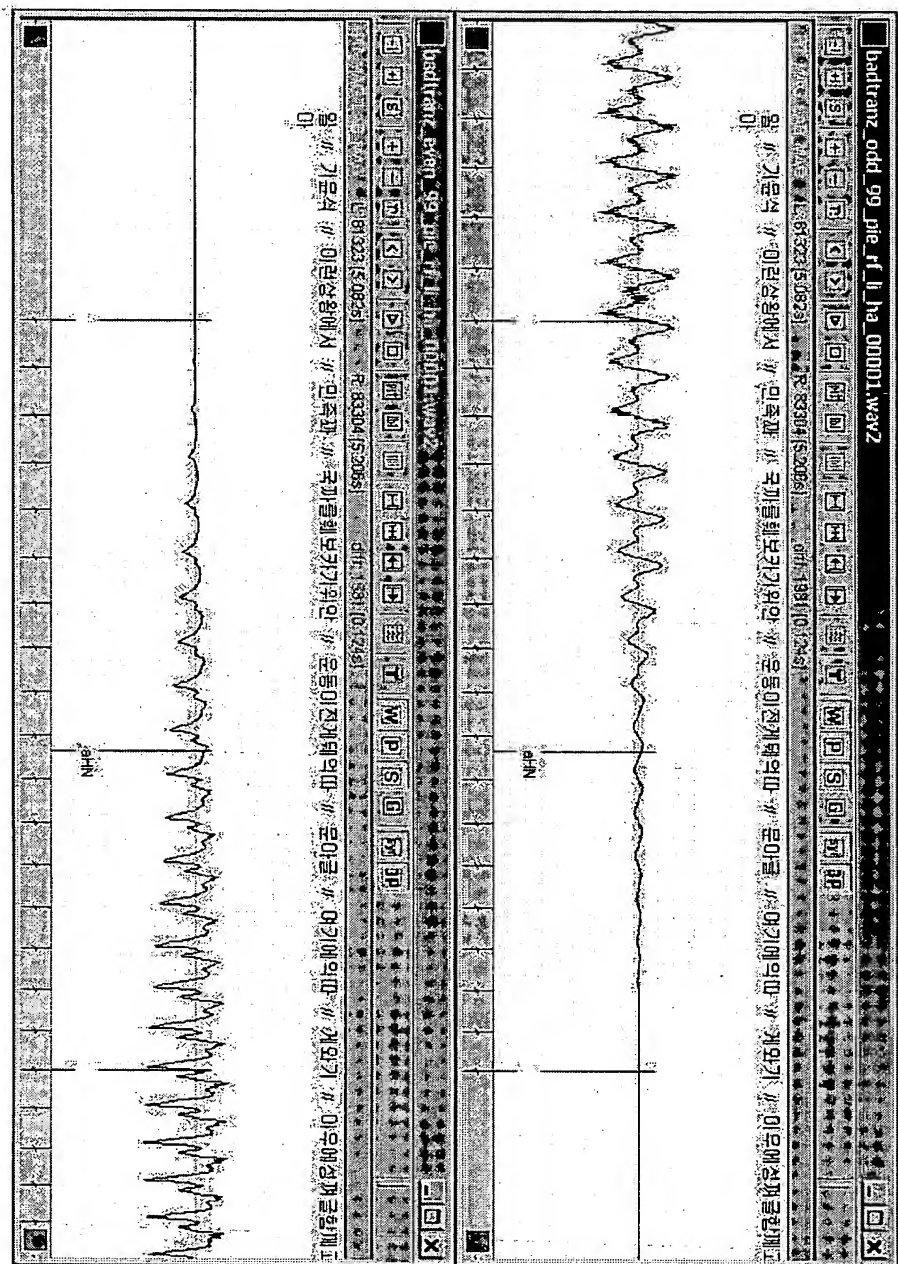


badtranz odd 99 pie II ha 00001.wav  
 101.2 bpm  
 NIP  
 bHN

badtranz even 99 pie II ha 00001.wav  
 101.2 bpm  
 NIP  
 bHN

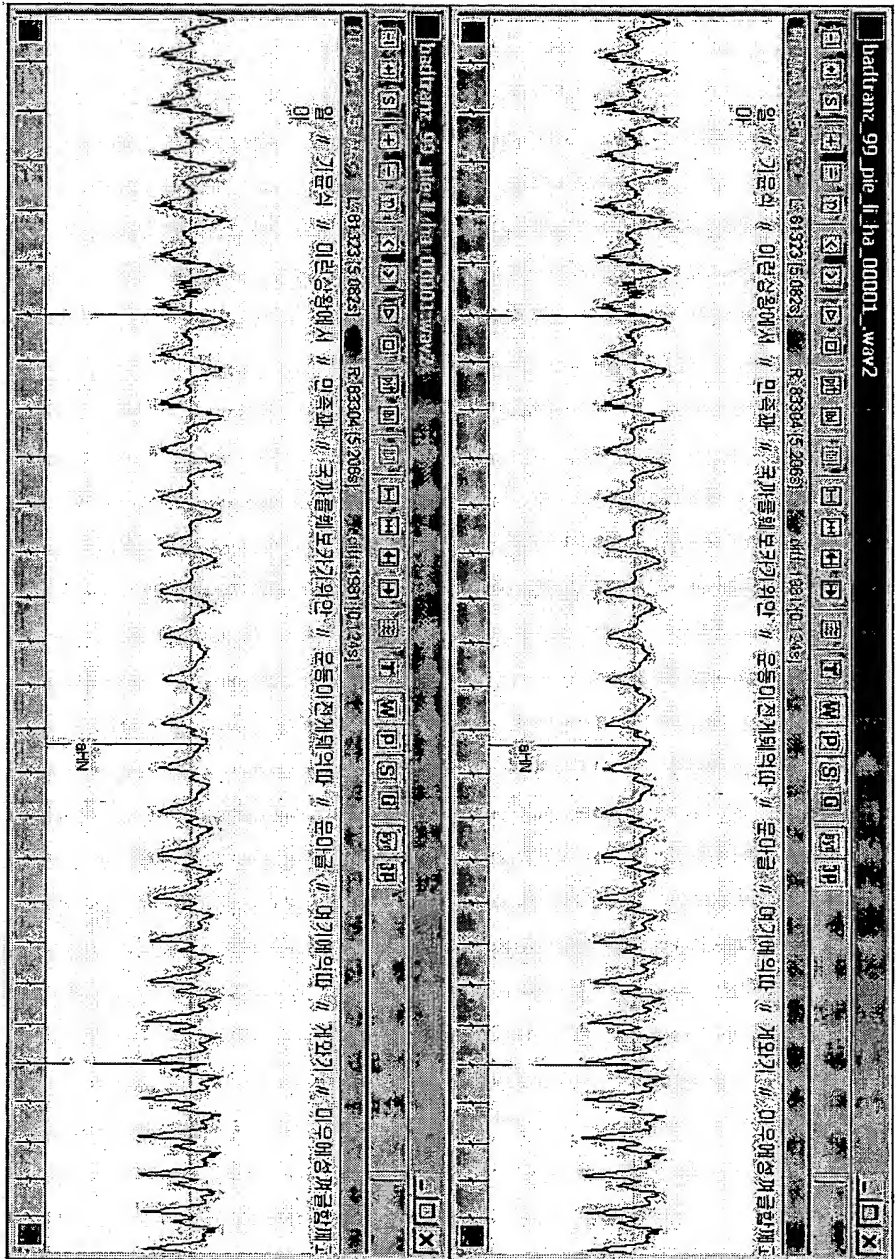


【도 11】

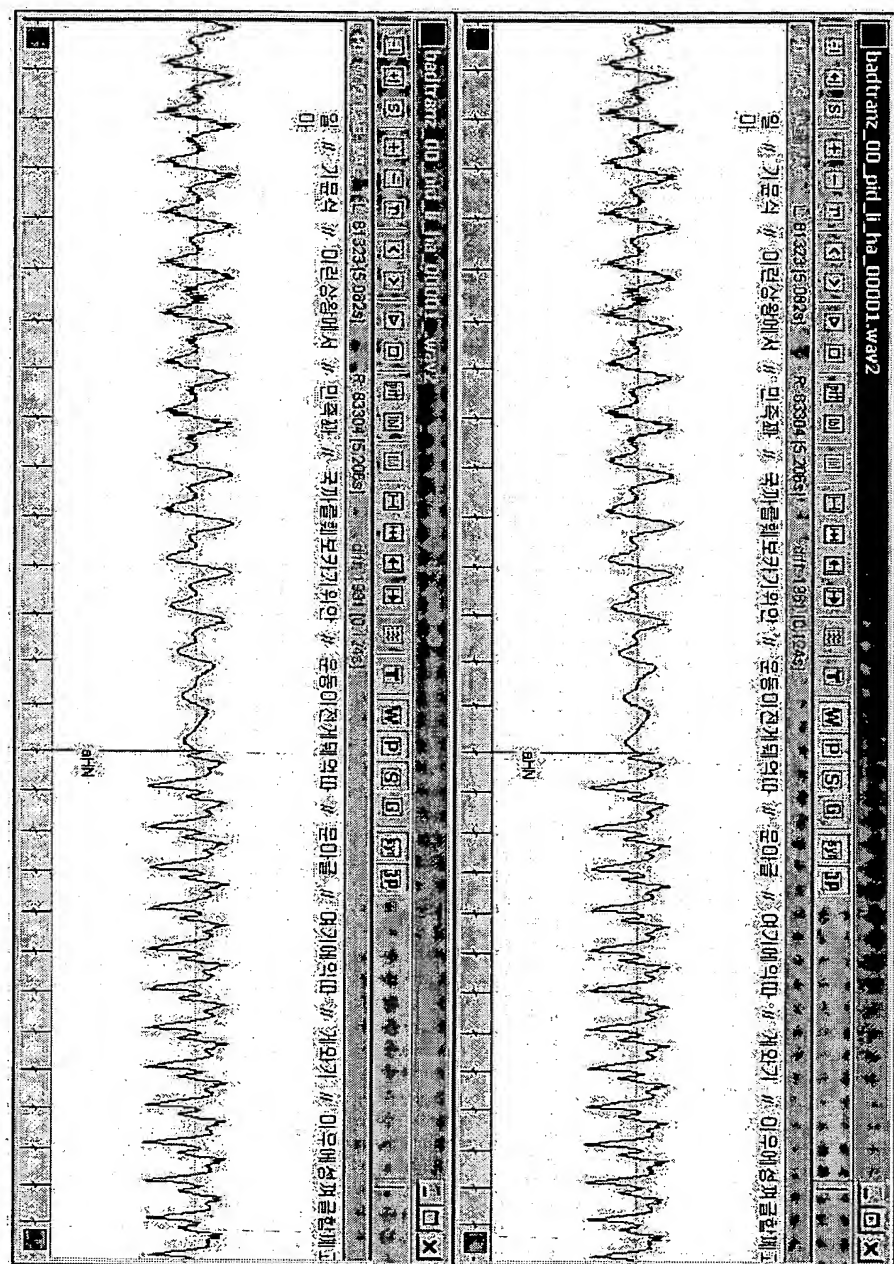




【도 12】



【도 13】



【도 14】

